

APPENDIX C.

Accuracy of the Data

CONTENTS

Confidentiality of the Data.....	C-1
Editing of Unacceptable Data.....	C-1
Sources of Error.....	C-1

CONFIDENTIALITY OF THE DATA

To maintain confidentiality required by law (Title 13, United States Code), the Bureau of the Census applies a confidentiality edit to assure published data do not disclose information about specific individuals, households, and housing units. The result is that a small amount of uncertainty is introduced into some of the census characteristics to prevent identification of specific individuals, households, or housing units. The edit is controlled so that the counts of total persons, totals by race and American Indian tribe, Hispanic origin, and age 18 years and over are not affected by the confidentiality edit and are published as collected. In addition, total counts for housing units by tenure are not affected by this edit.

The confidentiality edit is conducted by selecting a sample of census households from the 100-percent data internal census files and interchanging its data with other households that have identical characteristics on a set of selected key variables but are in different geographic locations within the same State. To provide more protection for "small areas," a higher sampling rate was used for these areas. The net result of this procedure is that the data user's ability to obtain census data, particularly for small areas and subpopulation groups, has been significantly enhanced.

EDITING OF UNACCEPTABLE DATA

The objective of the processing operation is to produce a set of data that describes the population as accurately and clearly as possible. To meet this objective, questionnaires were edited during field data collection operations for consistency, completeness, and acceptability. Questionnaires were also reviewed by census clerks for omissions, certain inconsistencies, and population coverage. For example, write-in entries such as "Don't know" or "NA" were considered unacceptable. For some district offices, the initial edit was automated; however, for the majority of the district offices, it was performed by clerks. As a result of this operation, a telephone or personal visit follow-up was made to obtain missing information. Potential coverage errors were included in the follow-up, as well as a sample of questionnaires with omissions or inconsistencies.

Subsequent to field operations, remaining incomplete or inconsistent information on the questionnaires was assigned using imputation procedures during the final automated edit of the collected data.

Allocations, or computer assignments of acceptable codes in place of unacceptable entries or blanks, are needed most often when an entry for a given item is lacking or when the information reported for a person or housing unit on that item is inconsistent with other information for that same person or housing unit. As in previous censuses, the general procedure for changing unacceptable entries was to assign an entry for a person or housing unit that was consistent with entries for persons or housing units with similar characteristics. The assignment of acceptable codes in place of blanks or unacceptable entries enhances the usefulness of the data.

Another way in which corrections were made during the computer editing process was through substitution; that is, the assignment of a full set of characteristics for a person or housing unit. When there was an indication that a housing unit was occupied, but the questionnaire contained no information for the people within the household, or the occupants were not listed on the questionnaire, a previously accepted household was selected as a substitute, and the full set of characteristics for the substitute was duplicated. The assignment of the full set of housing characteristics occurred when there was no housing information available. If the housing unit was determined to be occupied, the housing characteristics were assigned from a previously processed occupied unit. If the housing unit was vacant, the housing characteristics were assigned from a previously processed vacant unit.

SOURCES OF ERROR

In any large-scale statistical operation, such as the 1990 decennial census, human- and machine-related errors occur. These errors are commonly referred to as nonsampling errors. Such errors include not enumerating every household or every person in the population, not obtaining all required information from the respondents, obtaining incorrect or inconsistent information, and recording information incorrectly. In addition, errors can occur during the field review of the enumerators' work, during clerical handling of the census questionnaires, or during the electronic processing of the questionnaires.

To reduce various types of nonsampling errors, a number of techniques were implemented during the planning, development of the mailing address list, data collection, and data processing activities. Quality assurance methods were used throughout the data collection and processing phases of the census to improve the quality of the data. A reinterview program was designed to minimize the errors in the data collection phase for enumerator-filled questionnaires.

Several coverage improvement programs were implemented during the development of the census address list and census enumeration and processing to minimize undercoverage of the population and housing units. These programs were developed based on experience from the 1980 decennial census and results from the 1990 decennial census testing cycle. In developing and updating the census address list, the Census Bureau used a variety of specialized procedures in different parts of the country.

- For larger urban areas, the Census Bureau purchased and coded address lists, had the United States Postal Service (USPS) review and

update this list, and conducted a dependent canvass and update operation. Prior to mailout, local officials were given the opportunity to examine block counts of address listings (local review) and identify possible errors, and the USPS conducted a final review.

- For small cities and suburban and selected rural parts of the country, the Census Bureau created the address list through a listing operation that occurred in 1988 and 1989. For the addresses listed in 1988, the USPS reviewed and updated this list, and the Census Bureau reconciled USPS corrections through a field check; prior to mailout, the USPS conducted a final review of these addresses, and local officials participated in reviewing block counts of address listings.

Coverage improvement programs continued during and after mailout. The Census Bureau (rather than the USPS) delivered census questionnaires in the rural and seasonal housing areas listed in 1989 and in inner-city public housing developments. Computer and clerical edits and telephone and personal visit followups contributed to improved coverage.

If the 1990 census is not subject to count adjustment, the population counts shown in 100-percent data products will be tabulated from the entries for persons on all questionnaires. These counts will not be subject to sampling error. If count adjustment is done, a discussion of the count adjustment methodology and the appropriate methods for calculating sampling errors of adjusted counts can be found in appendix H. (Housing unit counts will not be subject to count adjustment.)